

Weekly Report

January 21, 2018

1 Work

1.1 降维

本周我们的降维程序已经完全把K-means聚类替换为四叉树和Fast Newman(社团检测算法)。目前算法从性能上提升了5-6倍的速度，在低维空间上使用KNN分类器的准确度有时会低0.3%左右，在CIFAR-100数据上面会低1%，我认为这可以通过后期调整一点参数修正。其实LargeVis和t-SNE比较的时候，由于准确率是做折线图的，所以即便是1%也很难看出来。我们目前是直接使用表格统计。

程序也查找了一遍内存泄漏等问题，总体上来说内存的开销已经比之前低了很多。

论文已经完成初稿，补充了很多实验的数据，基本已经成型。下周准备先给蔡登老师的学生付聪看一看。

1.2 图布局

下周可以开始图布局方法的实验，我认为总体思路可以分两步走：1) 改造已有降维算法。像t-SNE等会形成团状，这是不适用于图的，所以我们要做一点修正；2) 时间信息的考虑。简单的方法就是采用上一次迭代的结果，继续迭代，这样会减少两帧之间的误差。

1.3 工作进度

2 Paper Reading

深度学习的学习往往需要大量的数据，但是对于少量新的数据，神经网络无法快速学习，并且保持对新数据较高的泛化能力。为了解决这个问题，Meta Learning (Learning to Learn) 问题被提出。为了能够使得模型快速适应于新的数据和任务，Meta Learning基于以往的知识经验，只需要少量训练样本就能完成训练。

Table 1: 工作进度

TASK	PROGRESS	DATE
Graph Layout	下周分别进行两个实验	
dimension reduction	已经基本完成	
location2vec专利		
*2Vec survey		2.20

2.1 Meta-Learning with Memory-Augmented Neural Networks

文章将整一个训练数据看做一个序列，并且在LSTM的基础上又增加了一个External Memory用于存储已有的知识。当测试数据进入时，网络就会从External Memory找一些代表性的记忆，用于表达测试数据。

2.2 Matching Networks for One Shot Learning

首先，原有的模型可以对于新的数据抽取出特征表达。然后基于注意力机制，在这种特征表达的基础上，根据测试数据和支持集中特征表达的关系，进行预测。总的来说，是用注意力机制，将新的数据有已有的特征向量来表达。

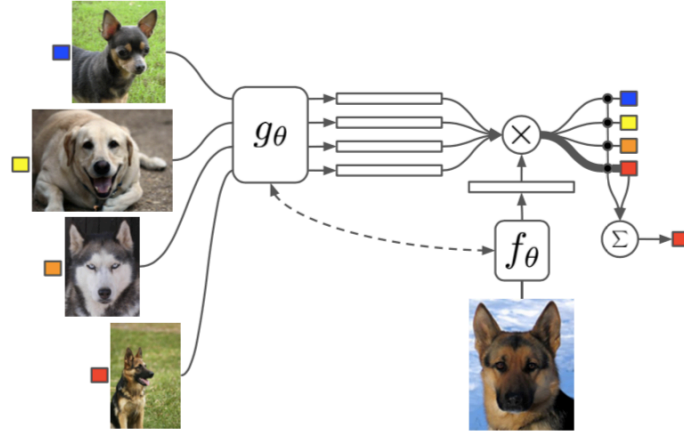


Figure 1: Matching Networks architecture

Figure 1: Matching Network

2.3 OPTIMIZATION AS A MODEL FOR FEW-SHOT LEARNING

文章将参数的更新与LSTM的学习过程联系起来，LSTM的状态就是需要学习的模型的参数。我认为用LSTM学习了一个CNN在遇到一个新数据下的梯度更新的过程，因此当遇到一个新的类别的数据时，LSTM (Meta-Learner) 就可以根据当前的Loss取得一个恰当的梯度。

2.4 Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks

这篇文章的角度更高一些，不是提出一个具体的方法，而是提出了一种更高层次优化模型的思想。文章首先是要训练一个模型，使得这个模型可以适用于不同的任务。这样的话，一方面是保证的模型可以有比较高的准确率，另一方面保证了模型不会局限于一个具体的问题，从而丧失了泛化能力。然后对于一个新的具体的任务，模型再一次通过简单的几次训练和梯度更新，就可以产生一个适用于新任务的模型。